



Coursework overview



Module assessment overview

Class test – 50%

You can do the demo in your lab in week 13 if you like

Coursework – 50%

- Large case study
- Many components are lab exercises
- You need to put it all together
- Add code for elements not yet covered
- Upload zip file of code to moodle by Friday 7 December 23:55pm
- Demo (15min) during week 14 or 15 (sign up for a slot in moodle)



Discussion of the Coursework Specification

See separate Word/pdf document for this



Coursework brief

- Test web page: <http://www.soc.napier.ac.uk/~40009856/CW/>
- Builds on lab exercises
 - *Which elements have you completed already?*
 - *Which elements are covered in remaining labs?*
- How to integrate?



Requirements - Marking criteria Part 1

1	Single web page download and parsing for links and outputting a summary of the total number of links, and a formatted list of the matched links [4 marks]	regex	webpage_get.py webpage_getlinks.py
2	Parsing and printing of unique email addresses and phone numbers [4 marks]		email_analysis.py
3	Parsing and printing of MD5 hashes found in the page source. Attempt to crack hashes found [4 marks]		dict_crack.py
4	Parsing and printing of any links to image files (e.g. gif,jpg,bmp) and documents (e.g. pdf,docx) files. [4 marks]		
5	Download the files into a directory created for this purpose. Report whether download successful or not. [4 marks]		urllib.request, urllib.parse os.path



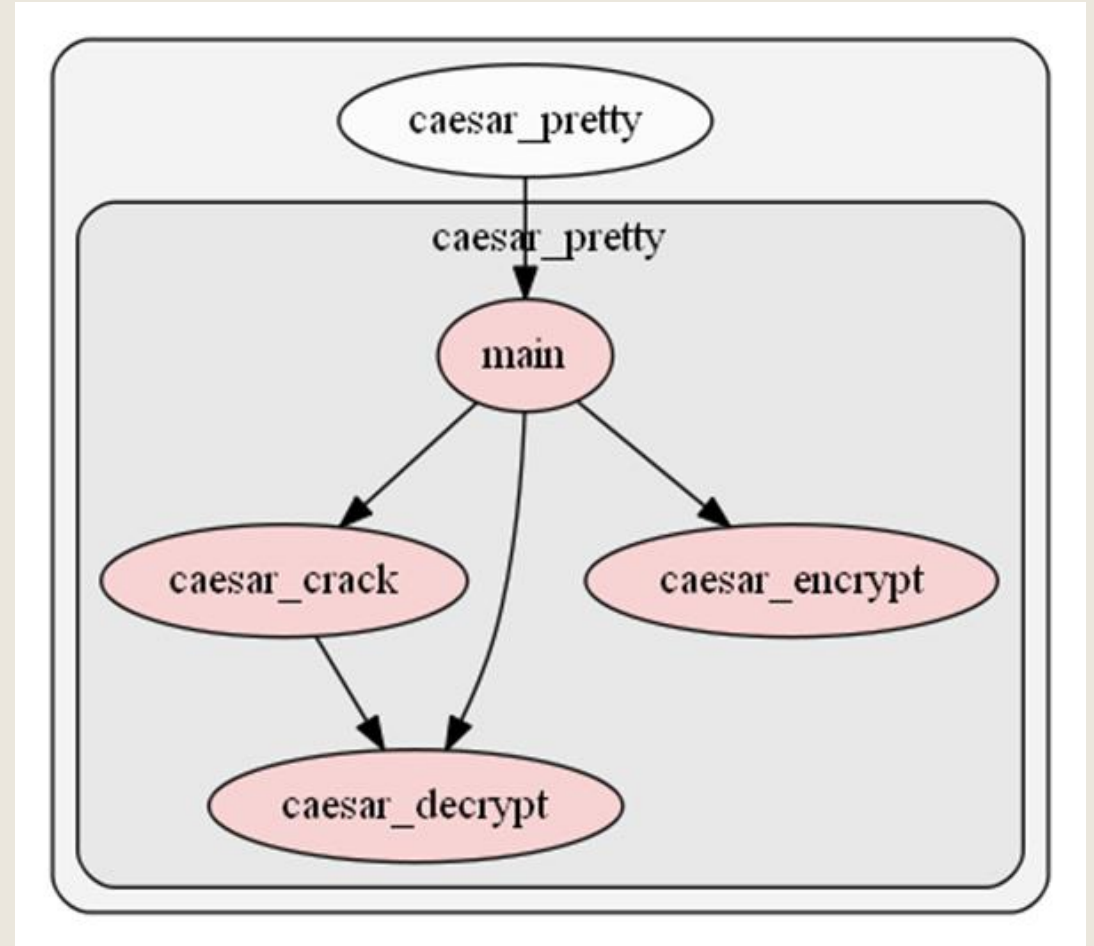
Requirements - Marking criteria

6	Files with the same name should be renamed and all downloaded. Check whether files with same name are duplicates or not. [4 marks]	os /os.path file_hash.py
7	Perform forensic analysis of the files downloaded: Report known bad files and file type / extension mismatches [4 marks]	file_hash.py file_type.py
8	Comprehensive Exception Handling. [4 marks]	
9	Good structure, clear dependency diagram, future-proof code. Good documentation (e.g. doc strings, comments, references) [4 marks]	Dependency diagram exercise
10	Explanations of code and answers to related questions [4 marks]	



What is a dependency diagram?

- Maps all the modules with their functions, and which function calls which (dependency), which module imports which
- Useful for code testing and debugging
- Useful for getting to know code written by others (including coursework submissions!)
- More dependencies (higher complexity) may reduce code efficiency and make it harder to maintain
- More dependencies may also have a good point, showing that code is modular.





Using code from others

Internet, friends, etc



Can I use code from my friends? Can I give my code to others?

- NO!!
- Exchanging code is Academic Deceit
- Penalties can be severe





- So does that mean we're not allowed to talk to each other about the coursework?
- You are allowed to talk ...
- You are allowed to help each other, e.g. with fixing errors ...
- ... as long as you don't give each other code!!



Can I use code from the internet?

- Yes, but:
 - *You need to clearly indicate where you have done this and give credit to your source*
 - *You need to make sure you fully understand it*
 - *Your code needs to run on Python3 in Windows - some websites will show examples for Python2 or linux*
 - *Don't copy an entire solution for this problem - I need to give YOU a mark, not the author of the code*



How to use code from the internet?

You need to give a reference (citation) that is specific enough

- for someone else to find it
- And to give credit

Specific author

Full URL

<https://stackoverflow.com/questions/10086572/ip-address-validation-in-python-using-regex>

Personal Napier Linux Python research Tools Hadoop NoSQL DigFor InfoSec DBForensics ProjectMgt hon's_projects DB T&L resources Digital skills from the

stackoverflow Questions Developer Jobs Tags Users Search...

```
return False # one of the 'parts' not convertible to integer
except (AttributeError, TypeError):
    return False # 'ip' isn't even a string
```

share improve this answer answered Apr 10 '12 at 11:59

Karl Knechtel 34.3k 4 54 87

This is better in my case. Thank you. – Khamidulla Jul 27 '15 at 5:48

add a comment

0

The following will check whether an IP is valid or not: If the IP is within 0.0.0.0 to 255.255.255.255, then the output will be true, otherwise it will be false:

```
[0<=int(x)<256 for x in re.split('\.',re.match(r'^\d+\.\d+\.\d+\.\d+$',your_ip).group(0))].c
```

Example: Code you're using

```
your_ip = "10.10.10.10"
[0<=int(x)<256 for x in re.split('\.',re.match(r'^\d+\.\d+\.\d+\.\d+$',your_ip).group(0))].c
```

Output:

```
>>> your_ip = "10.10.10.10"
>>> [0<=int(x)<256 for x in re.split('\.',re.match(r'^\d+\.\d+\.\d+\.\d+$',your_ip).group(0))].c
True
>>> your_ip = "10.10.10.256"
>>> [0<=int(x)<256 for x in re.split('\.',re.match(r'^\d+\.\d+\.\d+\.\d+$',your_ip).group(0))].c
False
>>>
```

share improve this answer edited May 10 '17 at 10:42 answered Mar 11 '16 at 12:55

Ranga 1 2

Meaning of the verb "tease" in context

How would a society isolated by monster attacks get food?

Extrude with exact measurements

How do you write a Stack Exchange answer?

How can tasks "refuse to freeze" on suspend?

Remove duplicates from a String

Do I need an RTC if my Raspberry Pi is continuously connected to the Internet?

What is Stream Sniping?

How do I get someone to call me by the name I prefer?

How to avoid accidental trespassing while exploring rural US?

Four year old started calling me by my first name. Is this problematic?

Speed to travel from Jupiter to Earth in only 3 hours?

Do Germans use "madame" often

Conjugacy in compact Lie groups

List operation depending on condition

Data scientist vs machine learning engineer

Want to see a Magic Card Trick?

First paid gig as an amateur; do I need insurance?

Why do you need to study a subject at a university level to teach it at a high school?

Can the NULL macro actually be a nullptr?

Understanding INKEY\$ in ZX BASIC

How much more dangerous is an encounter with the PCs outnumbered and how do I balance around it?

How to use code from the internet?

```
# this is ok, as long as you understand the code and can explain it yourself
def findIP_a(text):
    '''uses regex from https://stackoverflow.com/questions/10086572/ip-address-validation-in-python-using-regex (Ranga, 11/3/16)'''
    result = [0<=int(x)<256 for x in re.split('\.',re.match(r'^\d+\.\d+\.\d+\.\d+$',text).group(0))].count(True)==4
    return result
```

```
# this is good - you have split up the code to make it more understandable, and given a full reference
# in this case say "modified from" or "adapted from"
def findIP_a(text):
    '''modified from https://stackoverflow.com/questions/10086572/ip-address-validation-in-python-using-regex (Ranga, 11/3/16)'''
    found = re.split('\.',re.match(r'^\d+\.\d+\.\d+\.\d+$',text).group(0))
    count=0
    for x in found:
        if 0<=int(x)<256:
            count+=1
    if count==4:
        return found
```

I have not tested this code! (don't be surprised if this example doesn't actually work)

How NOT to use code from the internet?

```
# this is not ok, as you are pretending you wrote this expression (you are stealing ideas)
def findIP_c(text):
    result = [0<=int(x)<256 for x in re.split('\.',re.match(r'^\d+\.\d+\.\d+\.\d+$',text).group(0))].count(True)==4
    return result
```

```
# this is not ok, as the reference does not give enough detail for me to find it
def findIP_b(text):
    '''uses regex from stackoverflow'''
    result = [0<=int(x)<256 for x in re.split('\.',re.match(r'^\d+\.\d+\.\d+\.\d+$',text).group(0))].count(True)==4
    return result
```

I have not tested this code! (don't be surprised if this example doesn't actually work)



- Remember to use this approach for ALL exercises
- And in ALL your modules
- Applies also to references and ideas, not just code

